((((("inverted index" and document and "skip" and posting))))
Published before December 2003
Terms used: inverted index document skip posting

Sort results by: relevance — Save results to a Binder

Display results: expanded form — Open results in a new window

Refine these results with Ad
Try this search in The ACM (

Results 1 - 6 of 6

1 Self-indexing inverted files for fast text retrieval

Alistair Moffat, Justin Zobel
October 1996 ACM Transactions on Information Systems (TOIS), Volume 14 Issue 4
Publisher: ACM

Full text available: Pdf (484.52 KB) Additional Information: full citation, abstract, references, cited by, index terms

Bibliometrics: Downloads (6 Weeks): 23, Downloads (12 Months): 227, Citation Count: 77

Query-processing costs on large text databases are dominated by the need to retrieve and scan the inverted list of each query term. Retrieval time for inverted lists can be greatly reduced by the use of compression, but this adds to the CPU time required. ...

2 Compression of inverted indexes For fast query evaluation

Falk Scholer, Hugh E. Williams, John Yiannis, Justin Zobel
August 2002 SIGIR '02: Proceedings of the 25th annual international ACM SIGIR conference on Research and development in information retrieval
Publisher: ACM

Full text available: Pdf (174.13 KB) Additional Information: full citation, abstract, references, cited by, index terms

Bibliometrics: Downloads (6 Weeks): 17, Downloads (12 Months): 152, Citation Count: 31

Compression reduces both the size of indexes and the time needed to evaluate queries. In this paper, we revisit the compression of inverted lists of document postings that store the position and frequency of indexed terms, considering two approaches ...

Keywords: index compression, integer coding, inverted indexes, retrieval efficiency

3 Collection statistics for fast duplicate document detection

Abdur Chowdhury, Ophir Frieder, David Grossman, Mary Catherine McCabe
April 2002 ACM Transactions on Information Systems (TOIS), Volume 20 Issue 2
Publisher: ACM

Full text available: Pdf (191.32 KB) Additional Information: full citation, abstract, references, cited by, index terms

Bibliometrics: Downloads (6 Weeks): 19, Downloads (12 Months): 207, Citation Count: 24

We present a new algorithm for duplicate document detection that uses collection statistics. We compare our approach with the state-of-the-art approach using multiple collections. These collections include a 30 MB 18,577 web document collection developed ...

4 Building a distributed full-text index for the web

Sergey Melink, Sriram Raghavan, Beverly Yang, Hector Garcia-Molina
July 2001 ACM Transactions on Information Systems (TOIS), Volume 19 Issue 3
Publisher: ACM

Full text available: Pdf (651.72 KB) Additional Information: full citation, abstract, references, cited by, index terms, review

Bibliometrics: Downloads (6 Weeks): 10, Downloads (12 Months): 129, Citation Count: 9

We identify crucial design issues in building a distributed inverted index for a large collection of Web pages. We introduce a novel pipelining technique for structuring the core index-building system that substantially reduces the index construction ...

Keywords: Distributed indexing, Embedded databases, Inverted files, Pipelining, Text retrieval

5 Efficient query evaluation using a two-level retrieval process

Andrei Z. Broder, David Carmel, Michael Herscovici, Aya Soffer, Jason Zien
November 2003 CIKM '03: Proceedings of the twelfth international conference on Information and knowledge management
Publisher: ACM

Full text available: Pdf (248.95 KB) Additional Information: full citation, abstract, references, cited by, index terms

Bibliometrics: Downloads (6 Weeks): 22, Downloads (12 Months): 172, Citation Count: 12

We present an efficient query evaluation method based on a two level approach: at the first level, our method iterates in parallel over query term postings and identifies candidate documents using an *approximate evaluation* taking into account ...

Keywords: WAND, document-at-a-time, efficient query evaluation

6 Searching the Web

Arvind Arasu, Junghoo Cho, Hector Garcia-Molina, Andreas Paepcke, Sriram Raghavan
August 2001 ACM Transactions on Internet Technology (TOIT), Volume 1 Issue 1
Publisher: ACM

Full text available: Pdf (319.98 KB) Additional Information: full citation, abstract, references, cited by, index terms, review

Bibliometrics: Downloads (6 Weeks): 41, Downloads (12 Months): 675, Citation Count: 63

We offer an overview of current Web search engine design. After introducing a generic search engine architecture, we examine each engine component in turn. We cover crawling, local Web page storage, indexing, and the use of link analysis for boosting ...

Keywords: HITS, PageRank, authorities, crawling, indexing, information retrieval, link analysis, search engine

Results 1 - 6 of 6